Comments, suggestions, and related references welcomed!

Subgroups close to normal subgroups

GEORGE M. BERGMAN* AND HENDRIK W. LENSTRA, JR.*

Department of Mathematics, University of California, Berkeley, CA 94720, U.S.A.

Let G be a group and H a subgroup. It is shown that the set of indices $\{[H:H\cap gHg^{-1}] \mid g\in G\}$ has a finite upper bound n if and only if there is a normal subgroup $N \trianglelefteq G$ which is commensurable with H; i.e., such that $[H:N\cap H]$ and $[N:N\cap H]$ are finite; moreover, the latter indices admit bounds depending only on n. If the bounded index hypothesis is assumed only for g running over some subgroup $K\leq G$, the conclusion holds with "normal" weakened to "normalized by K".

More detailed information is gotten under the assumption that $\{[H:H\cap gHg^{-1}] \mid g\in G\} = \{1,p\}$ for p prime. In particular, when $p=2$ there exists $N \trianglelefteq G$ such that either H has index 2 in N, or N has index 2 in H.

1. THE $\{1,p\}$ CASE

This section contains the results referred to in the second paragraph of the abstract. The main results of the first paragraph are obtained in §2, which may be read independently. In the last three sections we extend both sets of results to the "normalized by K" context, obtain some modified bounds, and note some examples.

We begin with the result that started this investigation.

THEOREM 1. *Let G be a group, and H a subgroup. Then $[H:H\cap gHg^{-1}]\leq 2$ for all $g\in G$ if and only if G has a normal subgroup N such that either* (a) *$H\leq N$ and $[N:H]\leq 2$, or* (b) *$N\leq H$ and* $[H:N]\leq 2$.

Proof. "If" is clear. We shall prove "only if".

Let X denote the set of conjugates of H. We observe that there are no proper inclusion relations among members of X, since if x properly contained gxg^{-1}, the latter would have index 2 in x, hence g^2xg^{-2} would have index 4 in x, a contradiction. Thus, if two members of X are distinct, their intersection has index 2 in each.

Let us now choose any $x\in X$ and define two equivalence relations, $[x]$ and $<x>$, on $X-\{x\}$, letting $y[x]z$ mean $y\cap x = z\cap x$, and $y<x>z$ mean $yx = zx$ (equality of subsets). We claim

*This work was done while the authors were partly supported by NSF contracts DMS 85-02330 and DMS 87-06176 respectively. Electronic mail addresses of the authors at **cartan.Berkeley.Edu** or **ucbcarta.bitnet**: Bergman: **gbergman**, Lenstra: **hwl**.

AMS subject classifications (1985): Primary, 20E07; secondary, 20A15, 20B99.

(1) If $x, y, z \in X$ are pairwise distinct, then either $y[x]z$, or $y<x>z$.

Indeed, suppose first that $y \cap z \subseteq x$. Then $y \cap z \subseteq y \cap x$, and since both sides are subgroups of index 2 in y, we must have $y \cap z = y \cap x$. Likewise $y \cap z = x \cap z$; hence $y \cap x = x \cap z$, i.e. $y[x]z$. On the other hand, suppose $y \cap z \not\subseteq x$. Then $yx \cap zx$ consists of more than one left coset of x. But since yx and zx each consist of just two such cosets, we must have $yx = zx$, that is, $y<x>z$.

Now, it is straightforward to verify that if two equivalence relations R and S on a set T satisfy $R \cup S = T \times T$ (cf. (1)), then one of R, S is *indiscrete*, i.e. equal to $T \times T$. If the relation $[x]$ is indiscrete, then all the intersections $x \cap y$ $(y \in X - \{x\})$ are equal, hence their common value, the intersection of all conjugates of x, is a normal subgroup N of index 2 in x (unless X is a singleton, in which case x itself is normal), and in this case we have statement (b). If on the other hand $<x>$ is indiscrete then (again excepting the case $X = \{x\}$) the common value of the products yx will be closed under left multiplication by all conjugates $y \neq x$ of x, and clearly also under conjugation by members of x, hence under left multiplication by x; so it is closed under left multiplication by the subgroup N all these groups generate. Therefore it equals this normal subgroup, and as it consists of two left cosets of x, we have statement (a). ∎

Theorem 1 is equivalent to the $p = 2$ case of the next result, the proof of which is an elaboration of the same idea; but for later convenience, it will be preferable to formulate this proof in terms of the action of G on the left G-set G/H rather than its action on the set of conjugates of H.

If G is a group, and X a left G-set, then for $x \in X$, G_x will denote the stabilizer $\{g \in G \mid gx = x\}$. It is easy to verify that for any three elements x, y, z of a left G-set X, one has

$$(2) \qquad (G_x)z \supseteq (G_y)z \iff (G_x)(G_z) \supseteq G_y \iff (G_z)(G_x) \supseteq G_y \iff (G_z)x \supseteq (G_y)x.$$

THEOREM 2. *Let* G *be a group,* H *a subgroup, and* p *a prime number, and suppose* $\{[H:H \cap gHg^{-1}] \mid g \in G\} = \{1, p\}$. *Then* G *has a normal subgroup* N *such that either* (a) $H \trianglelefteq N$ *and* $[N:H] = p$, *or* (b) $N \trianglelefteq H$ *and* H/N *is isomorphic to a transitive permutation group on* p *letters.*

Proof. We see as in the proof of Theorem 1 that there can be no proper inclusion among conjugates of H; hence any two conjugates of H either coincide, or their intersection has index p in each of them. Letting X denote the G-set G/H, this says that for any $x, y \in X$, either $G_x = G_y$, i.e. $(G_x)y$ and $(G_y)x$ are singletons, or $(G_x)y$ and $(G_y)x$ both have cardinality p.

Let us now fix $x \in X$, and define two equivalence relations $[x]$ and $<x>$ on the set of points of X which have stabilizers distinct from G_x. Namely, $y[x]z$ will mean that the subgroup of G_x fixing all points of $(G_x)y$ coincides with the subgroup of G_x fixing all points of $(G_x)z$, while $y<x>z$ will mean that $(G_y)x = (G_z)x$. We shall again prove that for all y and z one of these relations holds.

Suppose $y[x]z$ does not hold. Without loss of generality we can assume that the pointwise stabilizer of $(G_x)y$ in G_x does not fix all points of $(G_x)z$. Now this stabilizer is a normal subgroup of G_x, hence the orbits of its action on $(G_x)z$, which by the above assumption are not all singletons, have equal cardinalities. But $(G_x)z$ has prime cardinality p, so the indicated stabilizer must act transitively on $(G_x)z$. This implies that the orbit of z under the possibly larger group G_y contains this set: $(G_y)z \supseteq (G_x)z$. But $(G_y)z$ cannot have cardinality larger than p, so we have equality: $(G_y)z = (G_x)z$. In particular, we have the reverse of our previous inclusion: $(G_x)z \supseteq (G_y)z$. By (2), this is equivalent to $(G_z)x \supseteq (G_y)x$. Now by the observation of our initial paragraph, the fact that neither y nor z is fixed by G_x tells us that x is fixed neither by G_y nor by G_z, hence each of the above orbits has cardinality

p, so this inclusion is an equality, proving $y<x>z$.

Hence by the same observation on equivalence relations made in the proof of Theorem 1, either $[x]$ or $<x>$ must be the indiscrete equivalence relation on elements having different stabilizers from x. Let us note that the set of such elements is nonempty, because our hypothesis on indices implies that H is non-normal. Let us now assume x chosen so that $G_x = H$, and let y be an element with $G_y \neq G_x$.

If the relation $[x]$ is indiscrete, then the pointwise stabilizer subgroup in G_x of the p-element orbit $(G_x)y$ is the pointwise stabilizer in G of all of X, hence is a normal subgroup $N \trianglelefteq G$, and we see that we have conclusion (b).

Suppose, then, that the relation $<x>$ is indiscrete. In this case $(G_y)x$ is the orbit of x under G_z whenever this stabilizer group is different from G_x. This constitutes a characterization of the set $S = (G_y)x$ in terms of the point x alone, hence S must be invariant under the action of G_x. Thus it is invariant under the actions of the stabilizers of all points of X, hence under the subgroup N that they generate, which will be a normal subgroup of G, in which $H = G_x$ has index $\mathrm{card}(S) = p$. Moreover the action of $H = G_x$ on the p-element set S cannot have any orbits of cardinality p, because it fixes $x \in S$, hence G_x must equal the pointwise stabilizer of this whole orbit of N, so it is a normal subgroup of N. This gives conclusion (a). ∎

Note that case (a) of the conclusion of the above Theorem implies the hypothesis of the Theorem, but case (b) does not. However, in §4 below we shall get more information about that case. We shall give examples of both cases in §5.

It is also possible to get results under weaker hypotheses, which have the above Theorem as special cases. Cheryl Praeger has proved such results under more general assumptions on the set of indices $\{[H:H \cap gHg^{-1}] \mid g \in G\}$, while I. M. Isaacs has obtained a result in which all assumptions on the indices, other than boundedness, are replaced by the condition that the $H \cap gHg^{-1}$ are maximal subgroups of H (personal communications).

2. Bounded Indices

We now turn to results holding when it is merely assumed that the set of indices $[H:H \cap gHg^{-1}]$ have a finite bound. This says that the images of H under the action of G by conjugation "stay close to" H; we will see that this happens if and only if H is close to a subgroup that is fixed under this action.

Recall that two subgroups H and K of a group G are called *commensurable* if both $[H:H \cap K]$ and $[K:H \cap K]$ are finite.

THEOREM 3. *Let G be a group and H a subgroup. Then the following conditions are equivalent:*

(i) *The set of indices $\{[H:H \cap gHg^{-1}] \mid g \in G\}$ has a finite upper bound n.*

(ii) *H is commensurable with a normal subgroup $N \trianglelefteq G$.*

Proof. (ii)⇒(i) is straightforward. To prove the converse, let n be as in (i), and let X be the left G-set G/H. We shall again write G_x for the stabilizer in G of an element $x \in X$. For any subset $S \subseteq X$, G_S will denote the pointwise stabilizer $\bigcap_{x \in S} G_x$.

Our hypothesis translates to say that for all $x \in G$, G_x acts on X with orbits of cardinality $\leq n$. Now for every finite nonempty subset $S \subseteq X$, let $m(S) \leq n$ denote the maximum of the cardinalities of the orbits of the action of G_S on X. Let m be the least value of the integers $m(S)$ as S ranges over all finite nonempty subsets of X, and let us call a finite nonempty set S with $m(S) = m$ a "strong" set. For S a strong set, let $N(S)$ denote the group of all elements of G that carry into themselves all cardinality-m orbits of G_S. Note that $G_S \leq N(S)$.

Clearly if S is a strong set, and T a finite set containing S, then T is also strong. In this situation, every G_T-orbit of cardinality m is a G_S-orbit; for if it were not, it would be contained in a strictly larger orbit of the group $G_S \geq G_T$, but by assumption G_S has no orbits of cardinality larger than m. It follows that $N(T) \geq N(S)$. Hence the groups $N(S)$ (S a strong subset of X) form an upward directed system of subgroups of G. Let N be the union of this system. Since this characterization of N in terms of X is translation-invariant, N is normal. (Concretely, this can be deduced by showing $gN(S)g^{-1} = N(gS)$.)

Now, taking a strong set S containing an element x such that $G_x = H$, we have

$$[H:H\cap N] \leq [H:H\cap N(S)] \leq [H:H\cap G_S] \leq n^{\operatorname{card}(S-\{x\})} < \infty,$$

where the next-to-last step uses our bound on indices of intersections, $[H:H\cap gHg^{-1}] \leq n$.

This gives half of commensurability; it remains to show that $[N:H\cap N] < \infty$; in other words, that the orbits of the action of N on X are finite. Clearly, if we can find a finite upper bound on the cardinalities of orbits of the groups $N(S)$, this will also bound the cardinalities of orbits of their directed union, N. Now every m-element orbit of G_S is an orbit of $N(S)$; taking any element x in such an orbit, we see that $N(S)$ meets exactly m left cosets of G_x. But G_x acts on X with orbits of cardinality $\leq n$, hence $N(S)$ acts on X with orbits of cardinality $\leq mn$, giving the needed bound. ∎

Curiously, knowing the above result, we can use it to get a better bound on $[N:H\cap N]$ than the one obtained in the proof. For observe that if S is a strong set, $G_S \leq N(S) \leq N$, so $N(S)$ is, like G_S and N, commensurable with H. Thus $[N:N(S)] < \infty$, hence there can be no infinite chain of groups between $N(S)$ and N. So, as N is the union of the directed system of groups of the form $N(S)$, it must equal one of these groups. Now we have noted that such a group has an orbit of cardinality m, but by normality of N, all its orbits have the same cardinality. Hence this cardinality is m; that is, $[N:H\cap N] = m \leq n$.

In contrast, the index $[H:H\cap N]$, for N constructed as in the above proof, cannot be bounded uniformly in terms of n. For example, let r be a positive integer, and let X be the set $Z_2 \times Z_r$, which we picture as a disjoint union of r 2-element sets. Let G_0 be the group of permutations of X which carry each of these subsets into itself, isomorphic as a group to $(Z_2)^r$, and let G be the group generated by G_0 and the cyclic shift of order r; i.e., the wreath product $Z_2 \operatorname{wr} Z_r$. Then X can be identified as a G-set with G/H, where H is the stabilizer of a single point of X, a subgroup of index 2 in G_0. Clearly, this stabilizer acts on X with orbits of cardinality ≤ 2. The least value of the function $m(S)$ defined in the proof of the Theorem 3 is 1, achieved when S meets all r copies of Z_2, so the subgroup N given by that proof is the trivial subgroup, whose index in H, 2^{r-1}, does not admit a uniform bound in terms of $n = 2$.

But in the case of the above example, we know from Theorem 1 that there is a normal subgroup N much closer to H. (Indeed, $N = G_0$ satisfies the conditions of that Theorem.) We shall now prove by a compactness argument that in the context of Theorem 3 one can always get an N (possibly different from the one constructed in the proof of that Theorem) satisfying a bound on $[H:H\cap N]$ uniform in n.

THEOREM 4. *For G, H, n as in condition* (i) *of Theorem 3, the N of condition* (ii) *can be chosen so that $[N:H\cap N] \leq n$ and $[H:H\cap N] \leq c(n)$ for some $c(n)$ depending only on n.*

Proof. Let n be fixed throughout the proof. We know that for any pair of groups G, H satisfying (i),

(3) There exists $N \trianglelefteq G$ such that $[N:H\cap N] \leq n$ and $[H:H\cap N] < \infty$.

We claim that such an N can be chosen so as to contain an intersection of finitely many conjugates of H. Indeed, given any N as in (3), let N_+ denote the intersection of all the conjugates of the group HN. Thus N_+ is a normal subgroup of G lying between N and HN. Hence $HN_+ = HN$. Now the two indices bounded in (3) can be written $[HN:H]$ and $[HN:N]$, so the corresponding expressions with N_+ in place of N also satisfy these bounds. In particular, since $[HN:N_+] = [HN_+:N_+] < \infty$ we see that the lattice of groups between N_+ and HN has finite length. Hence, as N_+ is the intersection of all the conjugates of HN, it can be written as an intersection of finitely many of these conjugates; hence it contains an intersection of finitely many conjugates of H, as claimed. Note that the latter finite intersection, being commensurable with H, will have finite index in N_+.

Now for any groups $G \geq H$ and positive integers i and j, let $P_{i,j}(G, H)$ denote the statement that there exists a family of $\leq i$ conjugates of H such that the union of some family of $\leq j$ right cosets of the intersection of these conjugates forms a normal subgroup N of G, with $[N:H\cap N] \leq n$. One verifies easily that each $P_{i,j}$ is equivalent to a first-order sentence about the pair (G, H), and the system of these sentences is directed under implication, since $P_{i,j}$ is weaker, the larger i and j are. We have just seen that every pair which satisfies (i) of Theorem 3 for our given n (a first-order condition) also satisfies some $P_{i,j}$. Hence by a standard application of the Compactness Theorem [1, Corollary V.5.6, p.213], there must be some $P_{i(n),\,j(n)}$ satisfied simultaneously by all such pairs of groups. (Concretely, if this were not so, an appropriate ultraproduct of counterexamples would satisfy no $P_{i,j}$.) Now for each (G,H), the N whose existence is asserted by $P_{i(n),\,j(n)}(G,H)$ satisfies $[H:H\cap N] \leq n^{i(n)}$, the desired bound. ∎

Such a compactness argument only gives the existence of a bound $c(n)$. Can we modify the proof of Theorem 3 so as to get an explicit bound? Let us return to the beginning of that proof, and for each integer $m \leq n$, let $h(m)$ be defined as the least cardinality of a finite nonempty subset $S \subseteq X$ such that all orbits of G_S have cardinality $\leq m$, or as ∞ if there is no such set S. Thus,

(4) $$1 = h(n) \leq h(n-1) \leq \ldots \leq h(1) \leq \infty.$$

In the proof of Theorem 3 we used the least m such that $h(m)$ was finite; but we have seen that this greatest finite value of $h(m)$ can be arbitrarily large. But if, instead, we look for the least $h(m)$ such that $h(m-1)$ is "much bigger" than $h(m)$ in some specific sense, then it should be possible to bound this value. So assume we have an m such that $h(m-1)$ is "much" larger than $h(m)$, in a sense to be specified as the need arises. Consider chains of subsets of X,

(5) $$S_1 \subseteq \ldots \subseteq S_r$$

such that (i) $m(S_i) = m$ for each i, (ii) $\operatorname{card}(S_1) = h(m)$, (iii) each S_{i+1} is of the form $S_i \cup g_i S_i$ for some $g_i \in G$, and (iv) the corresponding chain of groups $N(S_1) \leq \ldots \leq N(S_r)$ is strictly increasing. Note that $[N(S_r):H\cap N(S_r)] \leq mn \leq n^2$, and $[H:H\cap N(S_1)] \leq n^{h(m)}$, by the arguments used in the proof of Theorem 3. It follows that

(6) $$[N(S_r):N(S_1)] \leq n^{h(m)+2}.$$

This yields a bound on the length of chains (5), hence we can take a maximal such chain. Now if we try to adjoin to this chain another term, $S_{r+1} = S_r \cup g_r S_r$, we must either get $m(S_{r+1}) < m$ or $N(S_{r+1}) = N(S_r)$. By assuming $h(m-1)$ sufficiently large relative to $h(m)$, we can exclude the first alternative (details below); thus we can assume we have the second alternative for all choices of g_r. This means that for all $g \in G$, $N(S_r) = N(S_r \cup gS_r) \geq N(gS_r) = gN(S_r)g^{-1}$. Hence $N(S_r)$ is normal, and gives the desired subgroup.

Now, how big must $h(m-1)$ be compared with $h(m)$ to exclude the possibility $m(S_{r+1}) < m$ above? Observe that in a chain (5), each set S_{i+1} has at most twice the cardinality of S_i, while each group $N(S_{i+1})$ has at least twice the order of $N(S_i)$. Hence the ratio of the cardinalities of S_r and S_1 is bounded by (6), and we see that the condition $h(m-1) > 2n^{h(m)+2}h(m)$ will work. One can bound in terms of n the least value of $h(m)$ such that this condition is satisfied, and thus bound $[H:H \cap N(S_r)]$.

The resulting value for $c(n)$ (as in Theorem 4) is on the order of an n-times iterated exponential; we shall not work out a precise bound here (nor try to optimize every step of the above argument) because Peter Neumann has shown us a construction giving the much better bound $c(n) = n^{n^n}$ (personal communication, result to appear). Roughly, where we show that whenever $h(m)$ makes a large jump, the group we have called $N(S)$, which clearly satisfies $[N:H \cap N] \leq m$ and contains G_S, is normal, Neumann's proof shows that if $h(m)$ makes a more modest jump, the *normal* subgroup generated by G_S satisfies $[N:H \cap N] \leq n$.

3. K-NORMALIZING SUBGROUPS

We shall now give versions of the results of the preceding sections in which the conjugating elements are restricted to a subgroup $K \leq G$. We remark that if $H \leq K$, such results may be obtained as immediate consequences of the earlier results, by putting K in place of G, while they are trivial if $K \leq H$ (or even if K is contained in the normalizer of H); so the case of interest is that in which neither inclusion holds.

THEOREM 5. *Let G be a group, and H, K subgroups of G. Then*

(i) *$[H:H \cap kHk^{-1}] \leq 2$ for all $k \in K$ if and only if G has a subgroup N normalized by K, such that either $H \leq N$ and $[N:H] \leq 2$, or $N \leq H$ and $[H:N] \leq 2$.*

(ii) *If p is a prime number and $\{[H:H \cap kHk^{-1}] \mid k \in K\} = \{1, p\}$, then G has a subgroup N normalized by K, such that either $H \leq N$ and $[N:H] = p$, or $N \leq H$ and N contains the kernel of a transitive permutation representation of H on a set of p elements.*

(iii) *The set of indices $\{[H:H \cap kHk^{-1}] \mid k \in K\}$ has a finite upper bound n if and only if H is commensurable with a subgroup $N \leq G$ normalized by K.*

(iv) *In the implication $\Rightarrow$ of* (iii), *N can be chosen so that $[N:H \cap N] \leq n$, and $[H:H \cap N] \leq c(n)$ for some $c(n)$ depending only on n.*

Proof. Let $X = G/H$, and let $Y \subseteq X$ denote the image in X of $K \leq G$. Thus we now have not merely a transitive G-set, but one with a distinguished subset Y. Restrictions on the set of indices $\{[H:H \cap kHk^{-1}] \mid k \in K\}$ translate to the same restrictions on the set of cardinalities of orbits $(G_x)y$ where x and y both come from Y; but observe that these orbits themselves need not lie wholly in Y.

Happily, the greater part of the proofs of Theorems 1-4 can be adapted unchanged to this context if we take the elements named in these proofs, x, y etc., and likewise the named sets, S, T etc., to come from Y, but keep in mind the above observation about their orbits.

Here are details. We skip (i), since it will clearly follow from (ii). Under the hypothesis of (ii) we note, as in the proof of Theorem 2, that two elements of Y either have the same stabilizer, or else the image of each under the stabilizer of the other has cardinality p, and for $x \in Y$ we deduce as before that either the pointwise stabilizer subgroup in G_x of $(G_x)y$ is the same for all elements $y \in Y$ having stabilizers different from that of x, or the orbit $(G_y)x$ is the same for all such y. In the first case, the common subgroup of G_x is the kernel of the action of that group on the p-element set $(G_x)y$, and is contained in the pointwise stabilizer N of Y, a K-normalized subgroup of G_x. This N thus satisfies the second alternative conclusion of (ii). Now assume on the contrary that we have two elements y and z for which the above pointwise stabilizers of orbits are different. As in the proof of Theorem 2, we wish to show that $S = (G_y)x$ is invariant under the stabilizers of all points of Y. By what we have said, S is invariant under all such stabilizers that differ from G_x, but the argument by which we concluded in the proof of Theorem 2 that it was invariant under G_x no longer works – it only shows invariance under $G_x \cap K$. However, note that as in the proof of Theorem 2, we have $(G_x)z \subsetneq (G_y)z$, hence by (2) (with the roles of x and y reversed) we get $(G_y)(G_z) \supseteq G_x$, so since S is invariant under G_y and G_z, it is also invariant under G_x, as desired. We can now conclude as before that the group generated by the stabilizers of all elements of Y, which is normalized by K, contains G_x with index $\mathrm{card}(S) = p$. (We cannot say, however, that G_x acts trivially on $S = (G_y)x$, since it may move in orbits of cardinality $< p$ some points of S outside of Y; hence we cannot, as in Theorem 2, assert that G_x is normal in N.)

The proof of (iii) follows that of Theorem 3 exactly, subject only to the general changes noted above.

In proving (iv), the bound $[N:H\cap N] \leq m \leq n$ for the N constructed in the proof of (iii) is obtained as before. The next part of the argument needs modification, because N need not normalize H, hence HN need not be a group. Nevertheless, let us form the intersection of all conjugates of the set HN by members of K, calling this intersection M, and consider the group $N_+ = \{g \in G \mid Mg = M\}$. We see that N_+ is contained in HN and contains N, hence it satisfies $[N_+ : H\cap N_+] \leq [N:H\cap N] \leq n$ and $[H:H\cap N_+] \leq [H:H\cap N] < \infty$, and it is clearly normalized by K.

Now the set HN from which we started is a union of $\leq n$ right cosets of H. Thus we can formulate a family of first-order sentences $P'_j(G,H,K)$, saying that there exists a family of $\leq n$ right cosets of H such that, on taking the intersection M of all K-conjugates of their union, and forming the group $N = \{g \in G \mid Mg = M\}$, this satisfies $[H:H\cap N] \leq j$. (We don't need a separate condition bounding $[N:H\cap N]$, which will automatically be $\leq n$.) The Compactness Theorem again tells us that some $P'_{j(n)}$ must hold for all (G,H,K) satisfying the left hand side of (iii), and this completes the proof of (iv). One can also, as before, get an explicit bound by a more careful choice of m in the proof of (iii). ∎

4. Striking a Balance

In preceding sections we have gotten strong bounds on $[N:H\cap N]$, but very weak bounds on $[H:H\cap N]$. We shall now prove some results giving normal subgroups satisfying bounds of the reverse sorts. Our proofs will call on a couple of results in the group-theoretic literature; we are indebted to P. M. Neumann for pointing these out to us, and thus considerably shortening our arguments. The results of this section also rely on those of Sections 1 and 2.

Theorem 6. *Let G be a group and H a subgroup such that the set of indices $\{[H:H\cap gHg^{-1}] \mid g \in G\}$ has a finite upper bound n. Let N be a normal subgroup of G commensurable with H (which exists by Theorem 3). Then there exists a normal subgroup $M \trianglelefteq G$ containing N, such that $[H:H\cap M] \leq n$, and $[M:H\cap M] < \infty$ (equivalently, $[M:N] < \infty$).*

Proof. Dividing out by N, we are reduced to the case where N is trivial and H a finite subgroup of

G. In this situation let M be the least normal subgroup of G containing all elements of H which have only finitely many conjugates in G. Dietzmann's Lemma [3, §53] says that a finite union of finite conjugacy classes of elements of finite order in a group generates a finite subgroup; thus M is finite, which gives the final inequality. If we now divide out by M, we are reduced to showing that if H is a finite subgroup of a group G, such that all nonidentity elements of H have infinitely many conjugates, and H meets every conjugate of itself in a subgroup of index $\leq n$, then H must have order $\leq n$. To get this, we let G act by conjugation on the set of all conjugates of nonidentity elements of H, and apply [2, Theorem 1], which says that given a G-set X in which all elements have infinite orbits, and a finite subset $S \subseteq X$, there exists $g \in G$ such that $S \cap gS = \varnothing$. We conclude that H has a conjugate gHg^{-1} which has trivial intersection with H. The inequality $[H:H \cap gHg^{-1}] \leq n$ now gives the desired bound on the order of H. ▮

The above result, together with Theorem 4, which allows one to choose N so that $[N:H \cap N] \leq n$, mean, roughly, that the property of H of staying near its conjugates "splits" into three parts: The image of H in G/M is near its conjugates because it and they are groups of small order, $H \cap N$ is likewise near its conjugates because it and they have small index in the normal subgroup N, while the interesting part of the behavior, the proximity of $H \cap M/H \cap N \cong HN \cap M/N$ to its conjugates in G/N, is captured within the finite normal subgroup $M/N \trianglelefteq G/N$.

Under certain conditions, the numerical behavior of the indices $[H:H \cap gHg^{-1}]$ "splits up" in a quite precise fashion:

PROPOSITION 7. *Suppose, in the situation of Theorem 6, that* $N \leq H$. *Let* n *be the maximum of the indices* $[H:H \cap gHg^{-1}]$ $(g \in G)$. *Let* M *be constructed as in the proof of Theorem 6, let* $n_1 = [H:H \cap M]$, *and let* n_2 *be the maximum of the indices* $[H \cap M:(H \cap M) \cap g(H \cap M)g^{-1}] = [H \cap M:H \cap gHg^{-1} \cap M]$ $(g \in G)$. *Then* $n = n_1 n_2$.

Proof. For $g \in G$ maximizing $[H:H \cap gHg^{-1}]$, we have

$$n = [H:H \cap gHg^{-1}] \leq [H:H \cap M]\,[H \cap M:H \cap gHg^{-1} \cap M] \leq n_1 n_2.$$

To get the reverse inequality, let us now choose a $g \in G$ such that $[H \cap M:H \cap gHg^{-1} \cap M]$ assumes it maximum value, n_2. We recall that M/N is finite, while all nonidentity elements of the image of H in G/M have infinite orbits under the action of G induced by conjugation. Thus if we let G_0 denote the subgroup of G consisting of those elements whose action induced by conjugation fixes all elements of M/N, then G_0 has finite index in G, hence nonidentity elements of the image of H in G/M still have infinite orbits under this subgroup. Now applying the result from [2] quoted in the proof of Theorem 6 (or more conveniently in this case, [2, Theorem 2]), we can find an $h \in G_0$, conjugation by which takes the set of nonidentity elements of the image of gHg^{-1} in G/M to a set disjoint from the nonidentity elements of the image of H. It follows that

$$H \cap (hg)H(hg)^{-1} = H \cap (hg)H(hg)^{-1} \cap M = H \cap gHg^{-1} \cap M,$$

the last step because h acts trivially on M/N and we have assumed that H contains N. Now

$$n \geq [H:H \cap (hg)H(hg)^{-1}] = [H:H \cap M]\,[H \cap M:H \cap gHg^{-1} \cap M] = n_1 n_2. \quad ▮$$

COROLLARY 8. *Suppose* G *is a group,* H *a subgroup, and* p *a prime, such that* $\{[H:H\cap gHg^{-1}] \mid g\in G\} = \{1,p\}$. *Let* $N \trianglelefteq G$ *be a subgroup with the properties given in Theorem 2, and* M *the subgroup constructed using this* N *as in the proof of Theorem 6. Then exactly one of the following holds:*

(a) $H \trianglelefteq N = M$, *and* $[N:H] = p$,

(b) $N \le H \le M$,

(c) $N = M \trianglelefteq H$, *and* $[H:M] = p$.

Proof. If case (a) of Theorem 2 applies, then $H \trianglelefteq N$. It is easy to see that in this case the construction of Theorem 6 gives $M = N$. If case (b) of Theorem 2 applies, then $N \le H$, so we can apply Proposition 7, and conclude that either $n_1 = 1$, which gives (b) above, or $n_1 = p$, $n_2 = 1$. The last equation means that $H\cap M$ is normal in G; but the construction of Theorem 6 is such that M is the normal closure of $H\cap M$, so $M \le H$. This makes H/N an extension of M/N by the group H/M of order p. But by Theorem 2 H/N is representable by permutations of a p-element set, and such a permutation group cannot have a nontrivial normal subgroup of index p (consider the common cardinality of the orbits of such a subgroup); hence M/N must be trivial, so $M = N$, giving us case (c) above. ∎

Thus, H is either very close to a normal subgroup below it, or very close to a normal subgroup above it, or sandwiched between two normal subgroups a finite distance apart. (When $p = 2$ one of the first two of these statements in fact holds, by Theorem 1.)

Note that (b) is the only case of the Corollary in which H can act other than as a cyclic group of order p on its p-element orbits in $X = G/H$. This case is divided into two subcases by a well-known result of Burnside's [4, Theorem 7.3], which says that a transitive permutation group on a set of p elements is either doubly transitive, or is a subgroup of the $p(p-1)$-element group of affine transformations in one variable over the field of p elements. Let us now show that in the doubly transitive case we can, except when $p = 2$, exactly determine the index $[M:H]$. Indeed, we have a more general result:

PROPOSITION 9. *Let* G *be a group,* H *a subgroup, and* $n > 2$ *an integer such that* $\{[H:H\cap gHg^{-1}] \mid g\in G\} = \{1,n\}$. *Suppose that* H *acts doubly transitively on each of its n-element orbits in* $X = G/H$, *and let* $M \trianglelefteq G$ *be the least normal subgroup of* G *containing* H. *Then* $[M:H] = n+1$, *and* M *acts triply transitively on each of its orbits.*

Proof. Our hypothesis implies that every orbit of a stabilizer subgroup, $(G_x)y$ $(x,y\in X)$ has cardinality n or 1, and G_x acts doubly transitively thereon. By definition, M is the group generated by all the stabilizers G_x $(x\in X)$.

Let us define a "packet" to mean an $(n+1)$-element subset $P \subseteq X$ such that for each $y\in P$, the set $P-\{y\}$ is an orbit of G_y. We shall show that the packets are precisely the orbits of X under M, giving the first conclusion. Triple transitivity of M on each packet follows immediately.

We claim first that every n-element orbit of a stabilizer subgroup, $(G_x)y$, is contained in a packet. For by double transitivity of G_x on this orbit, we see that $G_x\cap G_y$ is transitive on $(G_x)y-\{y\}$, hence this set is contained in an n-element orbit of G_y, which we shall write $(G_x)y-\{y\}\cup\{z\}$. Let P be the $(n+1)$-element set $(G_x)y\cup\{z\}$.

By construction, $P-\{y\}$ is an orbit of G_y. Also, given distinct elements $u,v\in (G_x)y-\{y\}$, we observe that by double transitivity, some element of $G_x\cap G_u$ carries v to y, and some element of $G_y\cap G_u$ carries v to z. It is easily deduced that the n-element set $P-\{u\}$ is an orbit of G_u. Thus for every $w\in P$ except z we have shown $P-\{w\}$ to be an orbit of G_w. But it follows from the last case we proved that y and z are in the same orbit of a stabilizer G_u, and that P is invariant under this stabilizer, hence the property proved for y and P implies the same property for z and P, completing the proof that P is a packet.

We see from the definition that a packet is uniquely determined by any two of its elements; thus two packets that have at least two elements in common coincide. We shall now show that two packets cannot have just one element in common; i.e., that the packets partition X. This will imply that a packet must contain any orbit set $(G_x)y$ that it meets in even one point, and from this and the definition of a packet it follows immediately that the packets are the orbits of M.

So suppose two packets P and Q intersect in a singleton $\{x\}$. From double transitivity of G_x on $P-\{x\}$, it follows that the intersections $G_x \cap G_p$ as p runs over $P-\{x\}$ are all distinct, so clearly they cannot all be the same as all the intersections $G_x \cap G_q$ as q runs over $Q-\{x\}$. Say $G_x \cap G_p \neq G_x \cap G_q$. Thus the two sides of this inequality cannot both equal $G_p \cap G_q$; say $G_x \cap G_p \neq G_p \cap G_q$. The left-hand side is of index n in G_p (because P is a packet), and the right-hand side is of index at most n, hence the latter is not properly contained in the former, so we can find $g \in G_p \cap G_q$ not belonging to $G_x \cap G_p$. Since $g \in G_p$, this means $g \notin G_x$, i.e., g moves x. But as a member of G_p it must move x within the packet P, while as a member of G_q it must move x within the packet Q. Since P and Q are disjoint except for x, we have a contradiction, completing the proof of the Proposition. ∎

(The conclusion of the above result is false for $n = 2$, as shown by the example where G is a dihedral group and H a noncentral 2-element subgroup.)

We have not investigated systematically the non-doubly-transitive case of Corollary 8(b); but the only values of $[M:H \cap M]$ we have found in examples of this case (see next section) are $p+1$ and $2(p+1)$. We remark that in all cases of Corollary 8(b) that we know of, $[M:N] \leq (p+1)!$ and H has exactly $p+1$ distinct conjugates in G.

We leave to the interested reader further investigation of these questions, including the problems of finding explicit bounds on $[M:H \cap M]$ in the context of Theorem 6, of whether Proposition 7 can be extended to the situation where N is not necessarily contained in H (perhaps defining n_1 as in that Proposition, n_2 to be the maximum value of $[NH \cap M : NH \cap gNHg^{-1} \cap M]$ $(g \in G)$, and $n_3 = [N \cdot H \cap N]$), and of whether results like Theorem 6 can be proved in the "K-normalized" context.

5. Examples

We shall note here some relevant examples, leaving straightforward verifications of their properties to the reader.

We begin with examples illustrating the trichotomy described in Corollary 8 for $\{[H:H \cap gHg^{-1}] \mid g \in G\} = \{1, p\}$ (cf. sentence following the proof of that Corollary).

Example 1. H "very close to" a normal subgroup above it, but not to any normal subgroup below it: Let V be a vector space of large (possibly infinite) dimension over the field of p elements, let H be a subspace of codimension 1, let A be a group of automorphisms of V large enough so that the intersection in V of the orbit of H under A has "large" codimension, and let G be the semidirect product determined by the action of A on V. The hypothesis of Corollary 8 is easily verified; the "N" of conclusion (a) is the vector space V, in which H has index p. Any normal subgroup of G contained in H, on the other hand, has large index therein, by choice of A.

Example 2. H "very close to" a normal subgroup below it, but not to any normal subgroup above it: Let G be a simple group, infinite or of large finite order, having a cyclic subgroup H of order p. Here $N = \{e\}$.

Example 3. The "sandwiched" case. Let G be the symmetric group S_{p+1} and H the stabilizer of $1 \in X = \{1, \dots, p+1\}$. In this case, the N of Theorem 2 is the trivial subgroup and the M of Theorem 6 is all of G. Note that H is doubly transitive on its orbits, so Proposition 9 applies.

We can inflate this example by taking the direct product with the regular permutation representation of an arbitrary group F, so that G is the group $S_{p+1} \times F$ acting on $X = \{1, \dots, p+1\} \times F$, and H the stabilizer of $(1, e)$; in this version one has more than one "packet", and the stabilizer subgroup of a point has many orbits of order p and many orbits of order 1.

Note that in the above example, we could replace S_{p+1} by any doubly transitive subgroup thereof. If this group is triply transitive, then the stabilizers G_x of points of X are doubly transitive on their orbits. However, whether or not this is so, these examples satisfy $[M:H] = p+1$. The next example notes some cases where stabilizers act non-doubly-transitively and $[M:H] = 2(p+1)$.

Example 4. For the simplest case, replace $\{1, \dots, p+1\}$ and S_{p+1} in Example 3 (either the original or inflated version) by the 12-element vertex-set of an icosahedron, and its full symmetry group $A_5 \times Z_2$.

D. Goldschmidt has pointed out a family of examples with similar properties for an arbitrary prime $p \geq 5$: Let $G = \mathrm{PGL}(2, p) \times Z_2$, and let $H \leq G$ be the graph of the unique homomorphism from the "upper triangular" subgroup of $\mathrm{PGL}(2, p)$ (a semidirect product of Z_{p-1} and Z_p) onto Z_2.

Example 5. By taking a direct product of a case of Example 1, a case of Example 2, and a case of Example 3 or Example 4, we get a pair $H = H_1 \times H_2 \times H_3 \subseteq G = G_1 \times G_2 \times G_3$ such that for N, M as in Theorems 3 and 6, $H/H \cap M$, $H \cap M/H \cap N$ and $N/H \cap N$ are all nontrivial.

Turning to our "K-normalizing" results, observe that in Theorem 5 (ii), the two alternative inclusions both lack the normality condition of the corresponding inclusions in Theorem 2. The next two examples show that these normality conditions cannot always be attained.

Example 6. Let F be a nontrivial group, $p > 2$ a prime, $X = \{1, \dots, p\} \times F$, and G the group of permutations of X generated by $(S_p)^F$ acting on the F-tuple of copies of $\{1, \dots, p\}$, and F acting by translation on the second coordinates. (Note that in contrast to Example 3, the symmetric groups here act independently on the several p-element sets.) Thus, G is a wreath product of the symmetric group S_p by F. Let $Y = \{1\} \times F \subseteq X$ (a "section" of the projection to F), let H be the stabilizer of a point of Y, and let K be the group of elements of G carrying Y into itself. Although H has one $(p-1)$-element orbit in X, this orbit is disjoint from Y. All its other nontrivial orbits have cardinality p, so the hypothesis of Theorem 5 (ii) is satisfied. The only possibility for the N of that result is $(S_p)^F \times \{e\} \geq H$, and we see that H is not normal therein. (If we take F infinite, then N is the only K-normalized subgroup of G commensurable with H.)

Example 7. Let $p > 2$ be prime, and let X be the tree (nonempty connected acyclic graph) with p edges meeting at every vertex. (The tree with this property is unique up to isomorphism.) Let G be the full automorphism group of the tree X, let H be the stabilizer of a vertex $x \in X$, let $Y = \{x, y\}$ be an edge containing x, and let K be the group of all $g \in G$ carrying Y into itself. An N satisfying the conditions of Theorem 5 (ii) is given by the pointwise stabilizer of Y, which has index p in H and 2 in K. But no subgroup N that is normalized by H as well as K can be commensurable with H; for H and K together generate G, hence a subgroup normalized by both of them must be normal, and so must either be trivial or have all orbits infinite.

Note that in the last example, the set of stabilizer subgroups of elements of Y distinct from G_x is a singleton $\{G_y\}$, hence the relations $[x]$ and $\langle x \rangle$ used in the proof of Theorem 5 (ii) are both indiscrete;

yet only one of the two alternative conclusions of Theorem 5 (ii) holds. This shows that in the proof of that result, the use of the assumption of *non-indiscreteness* of the relation $[x]$ in obtaining the first alternative conclusion (a feature in which the proof differs from that of Theorem 2) is unavoidable.

However, the behavior illustrated by this example may be special to the case where there are exactly two distinct stabilizer subgroups of elements of Y (equivalently, where, writing E for the normalizer of H in G, we have $[K:K\cap E]=2$). Indeed, suppose there are >2 distinct stabilizers and that $[x]$ is indiscrete. Let us write $L(x)$ for the common value of the pointwise stabilizers of $(G_x)y$ in G_x, as y ranges over the elements of Y with $G_y \neq G_x$. We may now consider two subcases, according to whether $\langle x\rangle$ is or is not indiscrete. The case which formally most resembles Example 7 is that in which it is indiscrete. In this case, taking $x, y, z \in Y$ with distinct stabilizers, it is easy to verify that the pointwise stabilizers of $(G_y)x = (G_z)x$ in G_y and in G_z respectively coincide, i.e., that $L(y) = L(z)$, and to deduce that this will be a normal subgroup of H normalized by K, contrary to the behavior of the above example. In the case where $\langle x\rangle$ is not indiscrete, on the other hand, we do not know whether $L(x)$ is independent of $x \in Y$; i.e. is normalized by K.

It is not hard to see that several of the above classes of examples can be modified to yield cases where $\{[H:H\cap gHg^{-1}] \mid g\in G\} = \{1, n\}$ for n composite. We record one case that is not so obvious, a variant of the example of Goldschmidt given in Example 4,

Example 8. Let $G = \mathrm{PGL}(2, p) \times Z_{p-1}$ (p a prime ≥ 5), identify the second factor with the commutator-factor group of the "upper triangular" subgroup of the first factor, and let $H \leq G$ be the graph of the canonical homomorphism from this upper triangular subgroup onto this factor group. One can show that, via an appropriate change of basis, any distinct conjugate of the upper triangular subgroup of $\mathrm{PGL}(2, p)$ can be assumed to be the lower triangular subgroup. By comparing the maps of these two subgroups into Z_{p-1}, one can deduce that $\{[H:H\cap gHg^{-1}] \mid g\in G\} = \{1, p(p-1)/2\}$.

REFERENCES

[1] P. M. Cohn, *Universal Algebra*, 2nd ed., Reidel, Dordrecht, 1981. MR **82j**:08001. (1st ed. MR **31** #224 and **32** p. 1754.)

[2] B. J. Birch, R. G. Burns, Sheila Oates Macdonald and Peter M. Neumann, On the orbit-sizes of permutation groups containing elements separating finite subsets, *Bull. Australian Math. Soc.* **14** (1976) 7-10. MR **53** #5708.

[3] A. G. Kurosh, *Theory of groups*, Chelsea (1956). MR **15** p.501, **17** p.124, **18** p.188 (cf. also **9** p.267, **15** p.681, **22** #727, **40** #2740, **42** #1880, **50** #2314).

[4] Donald S. Passman, *Permutation Groups*, Benjamin (1968). MR **38** #5908.